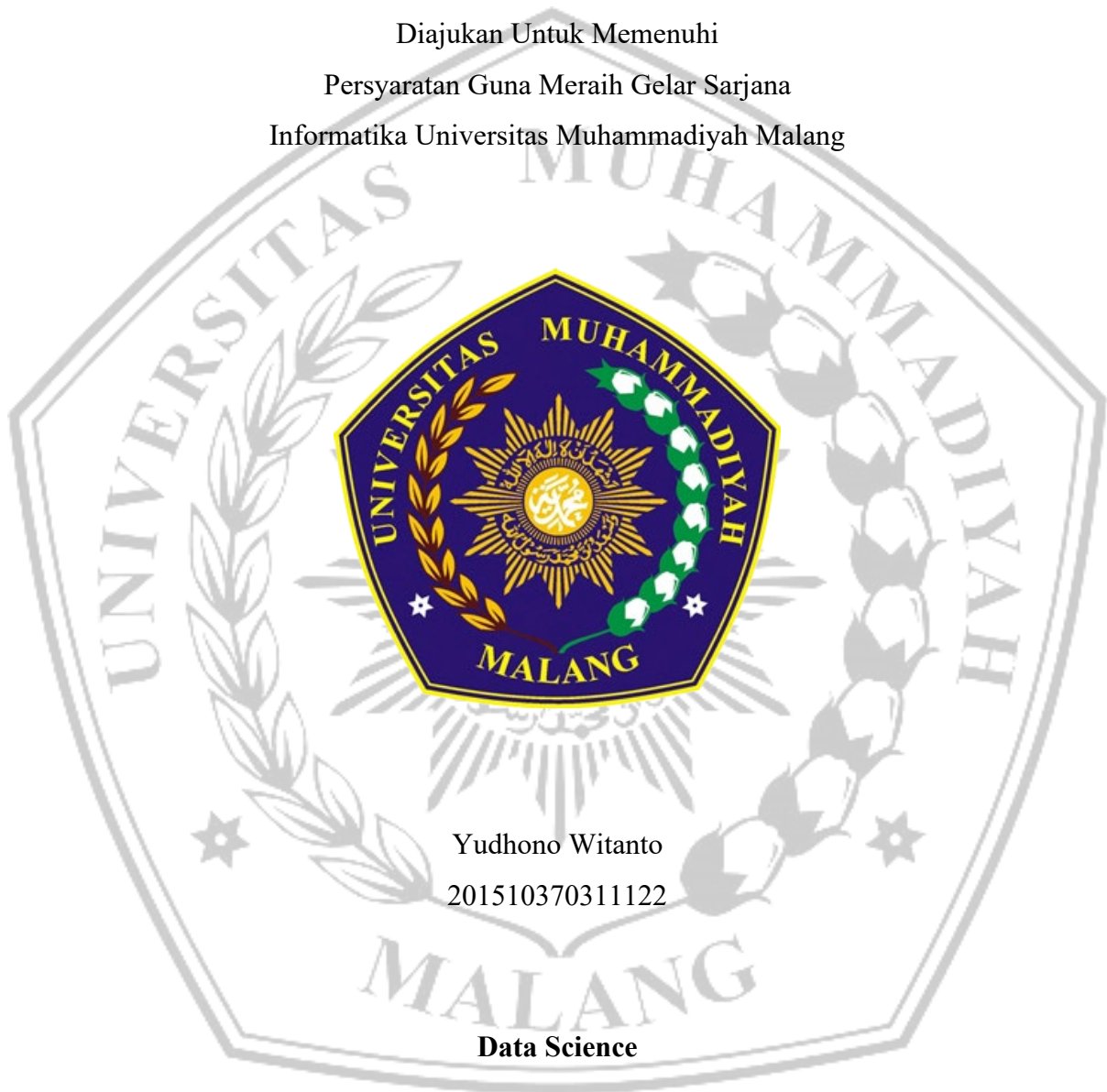


Optimisasi Algoritma *K-Means* Menggunakan *Artificial Bee Colony* pada *Content-Based Image Retrieval*

Laporan Tugas Akhir

Diajukan Untuk Memenuhi
Persyaratan Guna Meraih Gelar Sarjana
Informatika Universitas Muhammadiyah Malang



Yudhono Witanto
201510370311122

Data Science

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG
2019**

LEMBAR PERSETUJUAN

Optimisasi Algoritma *K-Means* Menggunakan *Artificial Bee Colony* pada
Content-Based Image Retrieval

Yudhono Witanto

201510370311122

Telah Direkomendasikan Untuk Diajukan Sebagai
Judul Tugas Akhir Di
Program Studi Informatika Universitas Muhammadiyah Malang

Menyetujui.

Dosen I

Dosen II

AGUS EKO MINARNO, S.KOM, M.KOM.

NIP. 108.1410.0540

YUFIS AZHAR, S.KOM, M.KOM

NIP. 108.1410.0544

LEMBAR PENGESAHAN

**Optimisasi Algoritma *K-Means* Menggunakan *Artificial Bee Colony* pada
*Content-Based Image Retrieval***

TUGAS AKHIR

Sebagai persyaratan Guna Meraih Gelar Sarjana Strata I
Teknik Informatika Universitas Muhammadiyah Malang

Disusun Oleh:

Yudhono Witanto

201510370311122

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
Pada Tanggal 10 Oktober 2020

Menyetujui,

Penguji I,

Penguji II,

Nur Havatin, S.ST., M.Kom.

NIP. 108.0907.0476

Aminudin, S.Kom, M.Cs.

NIP.108.1703.0594

Mengetahui,

Ketua Jurusan Teknik Informatika

Gita Indah Marthasari, S.T., M.Kom.

NIP. 108.0611.0442

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini:

NAMA : YUDHONO WITANTO
NIM : 201510370311122
FAK/JUR : TEKNIK / INFORMATIKA

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul **“Optimisasi Algoritma K-Means Menggunakan Artificial Bee Colony pada Content-Based Image Retrieval”** beserta isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui
Dosen Pembimbing,



Agus Eko Minarno, S.Kom., M.Kom.
NIP. 108.1410.0540

Malang, 20 Juni 2019
Yang membuat pernyataan,




Yudhono Witanto

ABSTRAK

Content-based Image Retrieval (CBIR) merupakan implementasi dari teknik *computer vision* pada kasus *image retrieval* yang merupakan kasus pencarian gambar digital pada database yang sangat besar. Pada penelitian ini diperkenalkan metode *clustering* baru untuk sistem CBIR, metode yang digunakan merupakan kombinasi antara algoritma *Artificial Bee Colony* (ABC) dengan *K-Means*. Tiga fitur digunakan untuk mengekstraksi fitur – fitur yang dimiliki gambar, yaitu: *RGB Color Feature*, *Edge Feature*, dan *Grey Level Co-occurrence Matrix* (GLCM). Ketiga fitur tersebut tergabung dalam algoritma *Multi Texton Co-occurrence Descriptor* (MTCD).

K-means merupakan algoritma pengelompokan yang banyak digunakan pada kasus *clustering* suatu data, *K-means* banyak digunakan karena implementasinya yang sangat mudah. Meskipun demikian, algoritma ini memiliki kelemahan, salah satunya yakni dalam penentuan titik awal *centroid*.

Artificial Bee Colony (ABC) merupakan algoritma optimasi yang cara kerjanya mengadopsi dari cara koloni lebah dalam mencari makanan. Metode ABC diketahui dapat memecahkan permasalahan *local optimum*, yang pada umumnya terjadi pada penggunaan *K-means* karena kelemahannya dalam menentukan *centroid*.

Metode yang akan digunakan pada penelitian ini merupakan kombinasi dari algoritma ABC dengan *K-means* untuk mengelompokkan fitur – fitur yang telah diekstraksi yang diimplementasikan pada dataset Corel-10.000 dan Batik. Kombinasi dari kedua algoritma ini dapat menjadi solusi dalam permasalahan pengelompokan data atau *clustering* di ranah *data science*. Penelitian ini bertujuan untuk mengetahui seberapa efektif dan efisien penggunaan kombinasi algoritma ABC dan *K-means* dalam *clustering* pada fitur dataset Corel-10.000 dan Batik.

Kata kunci: *Image retrieval, Ekstraksi fitur, Clustering, MTCD, Artificial Bee Colony, K-means.*

ABSTRACT

Content-based Image Retrieval (CBIR) is an implementation from computer vision techniques to the image retrieval problem, that is the problem for searching digital images in large databases. This research proposed a new method for CBIR system, combining Artificial Bee Colony and K-means algorithm. Three features are used to retrieve the images. These features are extracted by MTCD algorithm which included RGB Color Feature, Edge Feature, and Grey Level Co-occurrence Matrix (GLCM).

K-means is a clustering algorithm that used in many clustering problems, this algorithm is widely used because it is fluently works in every cases. However, this algorithm has weaknesses, one of which is in the initialization of centroid.

Artificial Bee Colony (ABC) is an optimization algorithm that works by adopting bee colony behavior in finding food. This method is known to solve local optimum problems, which generally occur in the use of K-means because of its weakness in determining centroids.

This research proposed a new method in clustering that combining ABC and K-means to cluster features that extracted from Corel-10.000 and Batik dataset. Combining these two algorithms can be a solution for data clustering in field of data science. The goal of this research is to determine how effective and efficient the combination of ABC and K-means algorithm in clustering the Corel-10.000 and Batik dataset.

Keyword: *Image retrieval, Feature extraction, Clustering, MTCD, Artificial Bee Colony, K-means.*

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan kasih sayang dan rahmat-Nya kepada kita, sehingga pada akhirnya mampu untuk menyelesaikan tugas akhir yang berjudul “Optimisasi Algoritma K-Means Menggunakan Artificial Bee Colony pada Content-Based Image Retrieval”.

Tujuan dari penyusunan skripsi ini guna memenuhi salah satu syarat untuk bisa menyelesaikan Studi S1 Program Studi Teknik Informatika di Fakultas Teknik Universitas Muhammadiyah Malang.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna dikarenakan terbatasnya pengalaman dan pengetahuan yang dimiliki penulis. Oleh karena itu, penulis mengharapkan segala bentuk saran serta masukan bahkan kritik yang membangun dari berbagai pihak. Semoga skripsi ini dapat bermanfaat bagi para pembaca dan semua pihak khususnya dalam bidang manajemen pemasaran.

Malang, 20 Juni 2019

Penulis,

Yudhono Witanto



DAFTAR ISI

LEMBAR PERNYATAAN.....	iv
ABSTRAK.....	v
ABSTRACT.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xii
DAFTAR SOURCE CODE.....	xiv
DAFTAR NOTASI.....	ii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Cakupan Masalah	4
1.5 Metodologi	5
1.6 Sistematika Penulisan.....	7
BAB II LANDASAN TEORI.....	9
2.1 <i>Content-Based Image Retrieval</i>	9
2.2 <i>Feature</i> (fitur).....	10
2.3 <i>Gray Level Co-Occurrence Matrix</i>	10
2.4 <i>Multi Texton Co-Occurrence Descriptor</i>	10
2.4.1 Multi Texton Histogram	11
2.4.2 Gray level Co-Occurrence Matrix (GLCM)	13
2.5 Algoritma Clustering K-means	14
2.6 Algoritma Artificial Bee Colony.....	15
2.7 Precision dan Recall	21
BAB III METODE PENELITIAN	23
3.1 Studi literatur.....	23
3.2 Dataset.....	23
3.3 Desain dan perancangan sistem.....	24
3.3.1 Ekstraksi Fitur Data	26
3.3.2 Aplikasi Metode ABC+KM.....	26
3.4 Pengujian	32

BAB IV IMPLEMENTASI DAN PEMBAHASAN	34
4.1 Lingkungan Implementasi.....	34
4.2 Implementasi	34
4.2.1 Implementasi Tahap <i>Preprocessing</i>	34
4.2.2 Implementasi Tahap Inisialisasi Sumber Makanan	36
4.2.3 Implementasi Optimasi Inisialisasi Sumber Makanan.....	37
4.2.4 Implementasi ABCKM pada Fase <i>Employed Bee</i>	39
4.2.5 Implementasi ABC+KM pada Fase <i>Onlooker Bee</i>	40
4.2.6 Implementasi <i>ABC+KM</i> pada Fase <i>Scout Bee</i>	41
4.2.7 Implementasi <i>Image Retrieval</i> menggunakan Sumber Makanan Terbaik 42	
4.3 Pengujian.....	47
4.3.1 Lingkungan Pengujian	47
4.3.2 Skenario Uji Coba.....	47
4.3.3 Uji Kinerja	49
BAB V KESIMPULAN DAN SARAN	57
5.1 Kesimpulan.....	57
5.2 Saran.....	58
DAFTAR PUSTAKA	59
LAMPIRAN A.....	62
LAMPIRAN B.....	67

DAFTAR GAMBAR

Gambar 2.1. Grid dan Texton pada MTH.....	12
Gambar 2.2. Texton pada MTCD	12
Gambar 2.3. Perbedaan deteksi texton pada MTH dan MTCD.....	13
Gambar 2.4. Konstruksi GLCM. (a) matrix gambar sebenarnya,.....	13
Gambar 2.5 Pseudocode algoritma K-means.....	15
Gambar 2.6 Flowchart Artificial Bee Colony.....	17
Gambar 2.7 Pseudocode Algoritma ABC.....	20
Gambar 4.1 Query 27_2761.jpg.....	43
Gambar 4.2 Hasil Retrieval dari query 27_2761.jpg.....	44
Gambar 4.3 Query 30_3005.jpg.....	45
Gambar 4.4 Hasil retrieval dari query 30_3005.jpg.....	45
Gambar 4.5 Query 17_1717.jpg.....	46
Gambar 4.6 Hasil retrieval dari query 17_1717.jpg.....	46

(Halaman ini sengaja dikosongkan)



DAFTAR TABEL

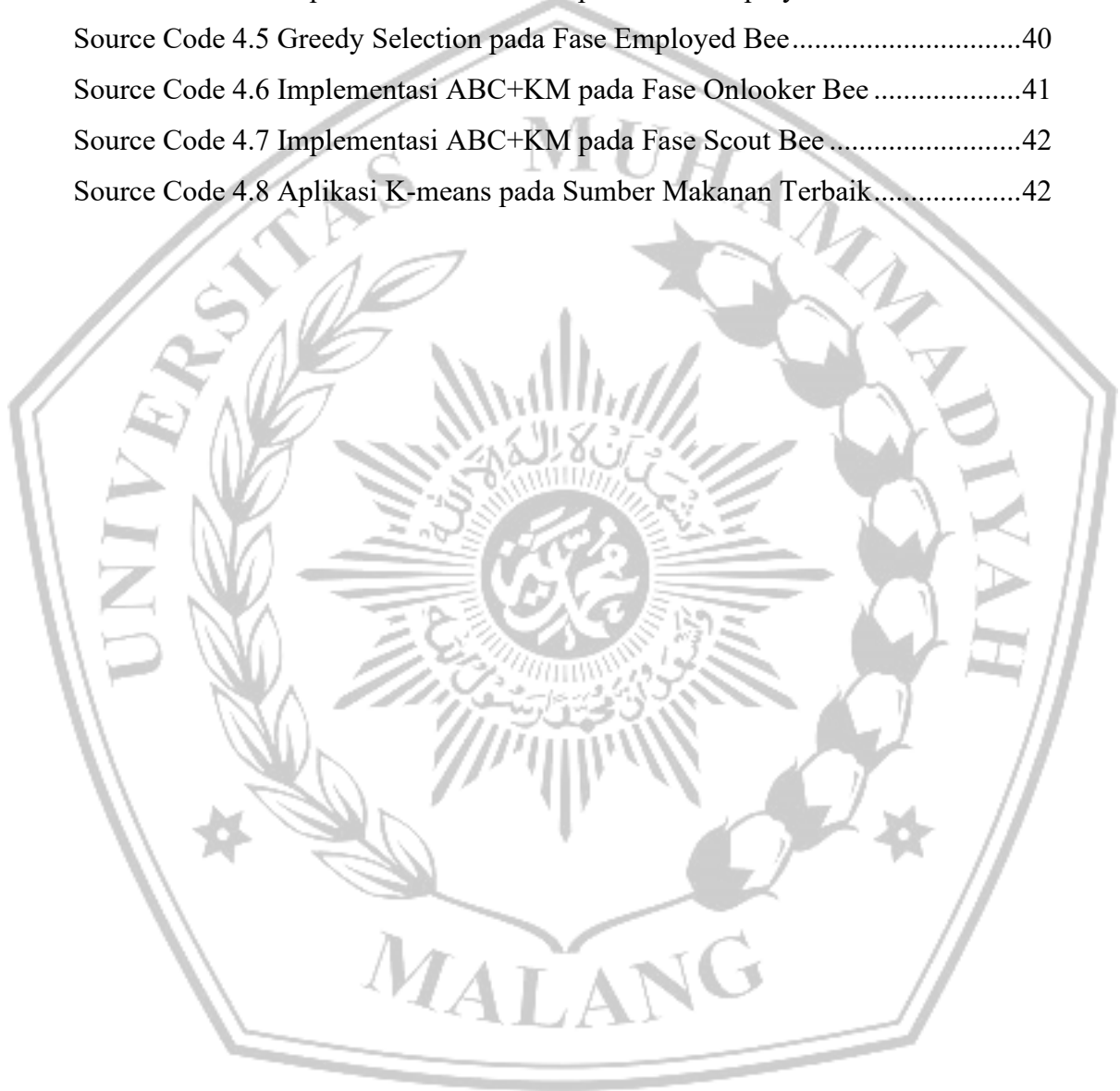
Tabel 4.2 Nilai precission dan recall.....	49
Tabel 4.3 Nilai Fitness Uji ABC+KM dataset Corel-10.000.....	50
Tabel 4.4 Nilai Fitness Uji ABC+KM dataset batik.....	51
Tabel 4.5 Nilai Precission Uji ABC + K-Means dataset Corel-10.000	52
Tabel 4.6 Nilai Recall Uji ABC + K-Means dataset Corel-10.000	52
Tabel 4.7 Nilai Precission Uji ABC + K-Means dataset Batik.....	53
Tabel 4.8 Nilai Recall Uji ABC + K-Means dataset Batik.....	53
Tabel 4.9 Perbandingan nilai performa untuk masing - masing metode	54
Tabel 4.10 Perbandingan nilai runtime untuk masing - masing metode.....	52
 Tabel Lampiran A 1 Sumber Makanan Terbaik Uji Coba dengan sumber makanan (SN) = 20, Iterasi KM = 25, percobaan 1, Dataset Corel-10.000	 62
Tabel Lampiran A 2 Sumber Makanan Terbaik Uji Coba dengan sumber makanan (SN) = 50, Iterasi KM = 25, percobaan 1, Dataset Corel-10.000	63
Tabel Lampiran A 3 Sumber Makanan Terbaik Uji Coba dengan sumber makanan (SN) = 50, Iterasi KM = 50, percobaan 1, Dataset Corel-10.000	64
Tabel Lampiran A 4 Sumber Makanan Terbaik Uji Coba dengan sumber makanan (SN) = 100, Iterasi KM = 100, percobaan 1, Dataset Corel-10.000	65
Tabel Lampiran A 5 Hasil Akhir Uji Coba dengan SN = 20, iterasi = 25	66
Tabel Lampiran A 6 Hasil Akhir Uji Coba dengan SN = 50, iterasi = 25	66
Tabel Lampiran A 7 Hasil Akhir Uji Coba dengan SN = 50, iterasi = 50	66
Tabel Lampiran A 8 Hasil Akhir Uji Coba dengan SN = 100, iterasi = 100	66
 Tabel Lampiran B 1 Data Ekstraksi Fitur Gambar menggunakan MTCD	 67
Tabel Lampiran B 2 Hasil Inisialisasi Sumber Makanan	68
Tabel Lampiran B 3 Hasil Optimasi Inisialisasi	72
Tabel Lampiran B 4 Hasil Fase Employed Bee untuk iterasi ke-1	77
Tabel Lampiran B 5 Hasil Fase Onlooker Bee untuk iterasi ke-1	86
Tabel Lampiran B 6 Hasil Akhir Sumber Makanan Terbaik	97

(Halaman ini sengaja dikosongkan)



DAFTAR SOURCE CODE

Source Code 4.1 Implementasi MTCD.....	35
Source Code 4.2 Inisialisasi Sumber Makanan.....	36
Source Code 4.3 Optimasi Inisialisasi Sumber Makanan.....	38
Source Code 4.4 Implementasi ABC+KM pada Fase Employed Bee.....	39
Source Code 4.5 Greedy Selection pada Fase Employed Bee.....	40
Source Code 4.6 Implementasi ABC+KM pada Fase Onlooker Bee	41
Source Code 4.7 Implementasi ABC+KM pada Fase Scout Bee	42
Source Code 4.8 Aplikasi K-means pada Sumber Makanan Terbaik.....	42



DAFTAR NOTASI

SN	Jumlah sumber makanan
D	Jumlah dimensi data
K	Jumlah <i>Cluster</i>
m_k	<i>Centroid</i>
C_k	<i>Cluster</i> ke- K
n_k	Jumlah pola data pada <i>cluster</i>
Z_p	Pola data
i	Merepresentasikan banyak $SN \{1, 2, 3, \dots, SN\}$
j	Merepresentasikan banyak dimensi $\{1, 2, 3, \dots, D\}$
u_i	Batas atas <i>solution space</i>
u_j	Batas bawah <i>solution space</i>
P_i	Nilai probabilitas
n_{ij}	<i>Food Source</i> baru
$\ Z_i - C_j\ ^2$	<i>Sum of Squared Error</i> (SSE) data Z_i ke <i>centroid</i> C_j
v_{ij}	Kandidat posisi sumber makanan baru berdasarkan yang lama
φ	Nilai acak dengan rentang $[-1, 1]$
P	Nilai probabilitas
E	Nilai <i>fitness</i>

DAFTAR PUSTAKA

- [1] Y. Liu, D. Zhang, G. Lu, and W. Ma, "A survey of content-based image retrieval with high-level semantics," vol. 40, pp. 262–282, 2007.
- [2] T. Quack, M. Ullrich, L. Thiele, and B. S. Manjunath, "Cortina : A System for Large-scale , Content-based Web Image Retrieval," 2004.
- [3] J. R. Smith and S. Chang, "VisualSEEk : a fully automated content-based image query system," 1996.
- [4] J. Ze, "SIMPLIcity : Semantics-sensitive Integrated Matching for Picture Libraries 1 Introduction," pp. 1–25.
- [5] U. M. Malang, "IMAGE RETRIEVAL USING MULTI TEXTON CO-," vol. 67, no. 1, pp. 103–110, 2014.
- [6] J. Rejito, A. S. Abdullah, and D. Setiana, "Image indexing using color histogram and k- means clustering for optimization CBIR in image database Image Indexing using Color Histogram and k -means Clustering for Optimization CBIR in Image Database," 2017.
- [7] "No Title," 2005.
- [8] D. Karaboga and C. Ozturk, "A novel clustering approach : Artificial Bee Colony (ABC) algorithm," vol. 11, pp. 652–657, 2011.
- [9] A. A. Amory, R. Sammouda, H. Mathkour, and R. M. Jomaa, "A Content Based Image Retrieval Using K-means Algorithm," pp. 221–225, 2012.
- [10] Y. Djenouri, A. Belhadi, and R. Belkebir, "Bees swarm optimization guided by data mining techniques for document information retrieval," vol. 94, pp. 126–136, 2018.
- [11] "jtap-039." .
- [12] G. Liu, L. Zhang, Y. Hou, Z. Li, and J. Yang, "Image retrieval based on multi-texton histogram," *Pattern Recognit.*, vol. 43, no. 7, pp. 2380–2389, 2010.
- [13] D. Zenzo, "[I," vol. 00, pp. 116–125, 1986.
- [14] G. Armano and M. R. Farmani, "Clustering Analysis with Combination of Artificial Bee Colony Algorithm and k -Means Technique," vol. 6, no. 2, pp. 141–145, 2014.

- [15] A. Bee, C. Abc, and T. Artificial, “Neural Networks.”
- [16] C. Chidambaram and H. S. Lopes, “A New Approach for Template Matching in Digital Images Using an Artificial Bee Colony Algorithm.”
- [17] P. Mansouri and B. Asady, “Solve Shortest Paths Problem by Using Artificial Bee Colony Algorithm Solve Shortest Paths Problem by Using Artificial Bee Colony Algorithm,” no. January 2014, 2016.





UNIVERSITAS MUHAMMADIYAH MALANG
FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA

Jl. Raya Tlogomas 246 Malang 65144 Telp. 0341 - 464318 Ext. 247, Fax. 0341 - 460782

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : YUDHONO WITANTO

NIM : 2015 1037 031 1122

Judul TA : Optimasi Algoritma *K-Means* menggunakan *Artificial Bee Colony* pada *Content-Based Image Retrieval*

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	10%
2.	Bab 2 – Tinjauan Pustaka	25 %	11%
3.	Bab 3 – Metode Penelitian	25 %	17%
4.	Bab 4 – Implementasi dan Pengujian	15 %	14%
5.	Bab 5 – Kesimpulan dan Saran	5 %	2%
6.	Makalah Tugas Akhir	20%	10%

